

385
982



BEER

FACULTY WORKING
PAPER NO. 982

THE UNIVERSITY OF ILLINOIS

DEC 5 1982

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Analysis of Audit Judgment Through an Expert System

Christopher Dungan

John S. Chandler

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 982

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign


November 1983

Analysis of Audit Judgment Through an Expert System

Christopher Dungan
University of South Florida (Sarasota)
John S. Chandler, Assistant Professor
Department of Accountancy

This research was supported by a grant from the Peat,
Marwick, Mitchell & Co., Research Opportunities in Auditing
Program.

This is a working paper and should not be quoted or reproduced
in whole or in part without the written consent of the authors.
Comments and suggestions should be forwarded to the authors.



Digitized by the Internet Archive
in 2011 with funding from
University of Illinois Urbana-Champaign

<http://www.archive.org/details/analysisofauditj982dung>

Abstract

The goal of the exploratory research described in this paper was to create an interactive computer program which would function as a dynamic, operating model of auditors' judgment and provide a tool for investigating the domain of audit judgment. The system would be operated in a noisy, natural setting to make credible audit decisions. The auditor's evaluation of the adequacy of the Allowance for Bad Debts constituted the functional area of interest in this study. The result of this work was the creation of AUDITOR--a model of an audit judgment in the form of an interactive computer program called an expert system. This paper describes how AUDITOR was designed and constructed. The results of two successful validation exercises are presented. A discussion of the insights into audit judgment provided by AUDITOR concludes this paper.

INTRODUCTION

Interest in the methods by which accountants make professional judgments has stimulated considerable research in recent years. Comprehensive reviews of this work may be found in Ashton (1982) and Libby (1981). Whether effected via regression modeling or focused on the probabilistic aspects of judgment, it is generally agreed that some of the research to date incorporates in its design, features which may operate as limitations on the external validity of the work. That is, typically, a relatively passive subject has been requested to perform a highly-structured and well-defined task via a simplified response scale in a laboratory setting while utilizing a limited set of data with which he has been provided.

Several approaches have been used to mitigate these limitations, for example, by introducing less-restricted search behavior. Biggs and Mock (1980) avoided the introduction of any restriction on subjects' cue choices by studying their spontaneous verbal protocols as they performed in an experimental setting. Abdel-Khalik and El-Sheshai (1980) allowed subjects to expand the number of factors which might be examined by choosing their own cues from among those provided in a shopping list. Shields (1980) allowed subjects to make choices of cues from information boards in a study of the effect of information load upon information search patterns.

Like the work on verbal protocols, this present study utilizes self-reports from the subjects themselves instead of inferring a decision model from the relationships between outputs and inputs. Although the validity of self-reports of mental processes is the subject of ongoing debate, c.f. Ericsson and Simon (1980), Nisbett and Wilson (1977), Payne, Braunstein, and Carroll (1978), Einhorn, Kleinmuntz, and Kleinmuntz (1979), certain aspects of the present study distinguish it from previous work.

The goal of this present, exploratory work was to create an interactive computer program which would function as a dynamic, operating model of auditor's judgment. The system would be operated in a noisy, natural setting to make creditable audit decisions. While operating within the confines of the available programming technology, the expert auditors who participated in building the system, themselves, freely determined the cues, their weights, and the form and size of the system. The result of this work was the creation of AUDITOR--a model of an audit judgment in the guise of an interactive computer program of a type properly identified as an expert system.

Auditors' evaluation of the adequacy of the Allowance for Bad Debts (ABD) constituted the functional area of interest in this study. This particular audit judgment was chosen for study and modeling for several reasons. It was one of the judgments most frequently named by the experts themselves when asked, "What critical audit judgment requires the attention of an experienced professional and is never entrusted to 'green' auditors?"¹

Also, when auditors make judgments of the adequacy of the ABD they are presented at a subsequent visit to the client's office with an opportunity for a referent outcome--the "outcome feedback" which is sometimes considered to be a necessary condition to sharpen and calibrate the expert's skills (Harrell, 1977; Ashton, 1981). That is, at a later date auditors can observe from the client's records which accounts have indeed proved to be uncollectible. Such a judgment seemed to the researchers more likely to provoke the consensus which Einhorn (1974) felt to be necessary requisite for expertise. Thus, in the absence of any other procedure for assuring that the decision process chosen for study would be one which would clearly be recognized as demonstrating expert judgment, the researchers believed expert judgment in auditing would most likely be fostered where the judgments were made under conditions which provided the judges with the possibility of objective outcome feedback. Finally, this area of judgment gave promise of tracability in initial interviews with the experts, in that they seemed able to recite the cues to which they attend when making the judgment, and the cues seemed similar among several experts.

EXPERT SYSTEMS

Expert systems are computer programs which offer consultative advice in a bounded knowledge domain on a level of competence often rivaling that of a human practitioner who is recognized as an expert in his field. Such systems are constructed by a system-builder utilizing the active collaboration of one or more of the experts themselves, who are challenged to justify the manner in which they make their judgments. These systems are now providing valuable consultative advice to clients in real-world settings (Michie, 1980).

Expert systems trace their lineage to attempts in the 1950s to create computerized problem-solving routines of broad generality, for example the Logic Theorist program of Newell, Simon, and Shaw (1958). This program and later generalized problem-solvers such as ARGUS (Reitman, 1965) and GPS (Ernst and Newell, 1969) made contributions to subsequent work but did not themselves display great breadth of achievement. More recent work proceeds from a belief that the high-level skill demonstrated by a human expert derives from his accumulated experience of a concentrated nature which enables him to perform optimally in situations to which that experience is pertinent but is of little help in genuinely novel situations (Simon, 1978).

Although frequently applied to programs of medical diagnosis or treatment, expert systems' technology appears to be appropriate wherever evidence of less than certain reliability must be evaluated by a skilled practitioner according to experientially based rules in order to make a judgment. Thus, the field of auditing bears similarities to domains already approached via the technology of expert systems. Comments of workers in the field of "knowledge engineering," as it is sometimes called, support this analogy.

The domain is one in which diverse factors must be identified and synthesized to form judgments, evaluate alternatives, and make decisions. Years of experience are brought to the problem at hand; experience and subjective judgment play a major role. The domain is not easily amenable to precise scientific formulation. (Duda et al., 1979)

The domain lacks a strong mathematical structure, is incorrigibly non-numerical, and is too complex for adequate analytical specification. (Michaelson, 1982)

The knowledge which the expert brings to the task is largely heuristic knowledge, experimental, uncertain--mostly good guesses and good practice in lieu of facts and rigor--much of this private to the expert. How else explain internships of guild-like apprenticeship to a presumed master of the craft? What the master really knows is not written in the textbooks. (Feigenbaum, 1979).

Examples of successful, currently operating expert systems are frequently reported in the popular press as instances of "artificial intelligence." For example, MYCIN infers disease identity in blood or meningitis infections and suggests antibiotic treatment (Shortliffe and Buchanan, 1975). PROSPECTOR evaluates core samples to infer the presence of significant mineral deposits (Duda et al. 1979). AL/X diagnoses causes of shutdowns occurring under the control of automatic safety devices on oil production rigs (Reiter, 1980).

AL/X (as did MYCIN) included in its design features the goal of separation of domain-specific knowledge from the computer control program which utilizes the knowledge. Thus, while AL/X was originally expert in only one realm of application its inference structure provides a foundation upon which can be built expert systems in other domains. AUDITOR utilizes the inference structure and control capabilities of AL/X, described in more detail in Appendix [A]. For a fuller treatment of its capabilities see Dungan (1983).²

AUDITOR FROM A USER'S PERSPECTIVE

The AUDITOR model was constructed in three distinct phases: initial modeling, refinement, and validation, all of which will be described in the course of this paper. However, certain aspects of this process of system-building can best be understood after the system has been described in use.

When a user operates the system interactively at a computer terminal for purposes of consulting AUDITOR for its advice concerning the large, delinquent receivable (which is the object of this inquiry), the

system asks him a series of questions about the extent of his personal knowledge of the presence or absence of the evidence cues required in the rules. AUDITOR begins each of its queries with the phrase, "How certain are you..."? which is prefaced to each of the rules in the rule base. For example, one query will be worded, "How certain are you that recent collections toward the delinquent portions of this account are proceeding satisfactorily?" The user's response provides a means for the system to acquire data about the underlying cue, suggested by the experts, which in this case concerns the debtor's recent payment performance.

The user responds to the system's query by typing a numeral called a certainty value (CV). CVs represent subjective certainty (or uncertainty) on a scale from -5 to 5. The response, "5," is given by the user of the system when he is certain that the matter referred to in the query is true. When he is certain the matter is false his response is "-5." In effect, the CV scale from .1 to 5.0 in absolute terms (unsigned) communicates increasing amounts of certainty, to which a positive sign is attached for truth and a minus sign for falsity. When the user has no information about the matter, he is unable to judge its truth or falsity, or he feels the question is irrelevant, he responds with "0." The researchers did not investigate individual differences in the use of this scale, although the process appears somewhat akin to the extraction of a utility curve, c.f. Newton (1977).

AUDITOR utilizes the data in the user's response in two ways. First, it updates the current strength of belief of the hypothesis, which in AUDITOR is stated in degrees of belief (DB), ($DB = 10 \cdot \log_{10} (\text{Odds})$). The nature of the updating depends upon the nature of the links between the rules and the hypothesis: Bayesian (for IF:THEN links), Logical (for AND, OR, NOT), or contextual (adaptable for unique linkages).

Second, the system inspects its rule base to determine which one of the remaining questions to ask next. It does this by following criteria built into the AL/X control (unless they have been modified by contextual links). The criteria are simply stated: Ask that question next which could possibly impact greatest upon the DB of the hypothesis, considering the inventory of questions, their diagnostic links, the current DB of the hypothesis, and the CV responses which might be given by the user. Thus, AUDITOR's inquiry proceeds in economical fashion down that line of questioning which could have the maximum impact upon the hypothesis.

This process of query, response, and update continues until (1) all questions in the system have been exhausted, (2) the user terminates the session, or (3) the updating process has caused the DBs of the hypothesis to reach a threshold level beyond which further questioning could cause no significant improvement, according to criteria adjustable within the system. In any of these cases, the system can then report its conclusion, which constitutes its "expert judgment" in the form of a statement of the current degrees of belief to which the hypothesis has been updated. In a typical consultation with AUDITOR this report might read:

The delinquent portion of this account should specifically be reserved-for in the allowance for bad debts to a substantial extent: RESERVE. Prior degree was 0.0, current degree is -19.5. At this point this goal is probability .05 or less.

This report informs the user of the system that the data provided by him has changed the degree of belief of the hypothesis from its initial state, 0.0 (.5 probability), to -19.5 (approximately .013 probability). At this point the user must decide for himself whether or not -19.5 DB (.013 probability) dictates to him a course of action, since there is insufficient experience with AUDITOR to provide guidance to relate these DB and probabilities definitively to decision points in an expert's judgment process.

CONSTRUCTION OF THE SYSTEM

A model of expert judgment, constructed in the form of an expert system, is assembled from components consisting of:

One or more hypotheses stating the judgment which is the end product of the expert's reasoning process.

Rules which express the relationship between the evidence (cues) and the hypothesis.

Parameters which express (1) the diagnostic value or impact which the expert believes are implied by each cue, and (2) the strength of the expert's beliefs prior to examination of any evidence, that is, the initial state of the system.

These components were assembled and tested during three stages: initial modeling, refinement, and validation. Initial modeling encompassed development of the hypothesis which is the goal of the system, interviews with the auditing experts to elicit cues, conversion of these cues into rules, and polling of the experts to elicit parameters adequate to construct a preliminary, yet operating, system. In the refinement stage, experts operated the model interactively on a computer terminal and presented their suggestions for improvements to be integrated into the system. These improvements consisted primarily of changes in parameters, additions of interactions between the rules, and a few instances of rewording of the cues. Finally, the completed system was exposed to validation by testing it to learn the extent to which a different group of experts would similarly acknowledge it to be expert in its performance. (Further references within this paper to "experts" will mean the several auditors who participated in the several stages of building, refining, and validating the AUDITOR system.)

INITIAL MODELING: HYPOTHESIS

Expert systems in other knowledge domains are constructed under a belief that an expert's information search and processing are focused upon and organized around hypotheses, one or more of which ultimately will express the judgment which he will render upon the data. There is speculation that the predecisional behavior of accountants and auditors is similarly organized around hypotheses (Libby, 1981). After selection of the judgment area for modeling the next step in the construction of AUDITOR consisted of formulating an hypotheses to reflect the expert's judgment. In an expert system this judgment is referred to as the goal hypothesis.

AUDITOR's single, goal hypothesis is expressed, "The delinquent portion of this account should specifically be reserved for in the allowance for bad debts to a substantial extent." (For ease of reference and recognition, the name of this hypothesis, RESERVE, as well as the name of each rule in the system will be written in capital letters.) The phrase, "this account" refers to the one individually large account--or in some cases a single invoice--which is under scrutiny by the auditor. "To a substantial extent" was agreed among the participants of this project to mean all cases in which they judged recovery is likely to be no more than an amount considered insignificant in relation to the delinquent balance.

RESERVE is intended to reflect the approach which the experts testified they visualized when scrutinizing the individually large, delinquent accounts of a commercial, audit client. It can be labelled a "worst-yet-possible-case" approach. This approach envisions that each, individually large account (or invoice) be judged collectible or uncollectible. The magnitude of the client's allowance for bad debts (ABD) is then considered adequate or inadequate in comparison to the aggregate of the large, delinquent accounts which have been judged uncollectible.⁴

DEVELOPMENT OF THE RULE BASE

The rule base for AUDITOR was developed from the cues provided by four practicing auditors chosen from among the eight who were consulted initially when choosing an area for modeling. Those actively participating in the bulk of the project were three managers and one senior on the audit staff of a CPA firm of international scope.⁵ (By comparison, for the development of expert systems in other fields the wholehearted cooperation of one practitioner who is clearly expert is often considered adequate.)

In individual interviews these auditing experts recalled and named the cues to which they said they attend when evaluating the collectibility of a client's delinquent receivables. For example, one cue

relates to whether the debtor continues to be an active customer. Other cues relate to the opinions of the client's credit manager as to collectibility of the delinquent account, response to confirmation requests, and contents of the client's credit file on the delinquent customer. A total of twenty-five such cues were obtained, differing widely among themselves in the diagnostic impact which they have upon the auditor's conclusion. The complete list may be examined in Dungan (1983).

After eliminating from the list obvious redundancies and a few errors in transcribing, the researchers then transformed each of the cues into a simple rule in IF:THEN form. For example, a cue named by one of the experts was: "The customer's stated intent regarding payment." This became the rule called NOTPAY: "The customer has stated his intent to pay little or nothing of the delinquent account," which is processed by AUDITOR utilizing a Bayesian revision as IF,NOTPAY: THEN (to the extent determined by parameters), RESERVE. The connector AND was later added (during the refinement stage of the work) in order to provide linkages between rules said by the experts to be interrelated.⁶ The connector NOT was used occasionally to achieve a more natural wording of a rule. (Although available for use through AL/X, the connector OR was not used.)

PARAMETERS

Next, parameters were developed for the rules--the diagnostic weights which express what the experts believe to be the evidential importance of each of the cues. This step began with a polling of the four auditing experts. Each of their rankings of "strong," "moderate," "weak," and "no effect" for each of the rules were translated into a quantitative scale on an expedient and preliminary basis. For example, all four experts rated the impact of NOTPAY as "strong." These preliminary values were later changed to the values contained in the final version of AUDITOR when the experts themselves operated the system and offered criticisms aiding in its refinement.

AL/X expresses the diagnostic impact of the evidence in the form of Positive Weights (PW) and Negative Weights (NW). As the AUDITOR system is operated interactively, the degree of belief attached to the goal hypothesis (RESERVE) is incremented by a value which is a function of the PW and the Certainty Value (CV) of the user's response if the user has keyed-in a positive CV in answer to the system's query, and upon the NW and the CV when the user's response is a CV in the negative range. The role of PW and NW are explained more fully in Appendix [A].

REFINEMENT

Armed with a basic set of rules and parameters the researchers assembled a working model of AUDITOR with the aid of the AL/X software package. For purposes of refinement, this working model was returned to the experts to secure their suggestions for improvements. As a

result of this refinement stage, major changes were made in the PW and NW, numerous interactions were added, and a few changes in the composition of the rule base were found to be desirable. An example of these changes, the rule called WORKOUT was originally developed from a cue which expressed the auditor's interest in a delinquent debtor's performance toward fulfilling a "workout alreement" or similar negotiated understanding between the client and the debtor. Before refinement the rule was stated, "Payments are being received currently under a workout agreement." As a result of reactions received during refinement of the system, this rule was restated to read, "Recent collections toward the delinquent portions of this account are proceeding satisfactorily," in order to eliminate the restrictive nature of the reference to "workout agreements."

Also at this stage of the work interactions in the form of AND statements were added to the rule base upon the suggestions of the experts. For example, the individual effects of the rules called LEGAL and NOTPAY are enhanced when both of these cues are found present at the same time. That is, when a delinquent debtor has stated his intent not to pay (NOTPAY) and the same debtor also presents a counterclaim which would appear to make legal action fruitless (LEGAL), the AUDITOR system operates by implementing each of these rules individually and also implements a third rule, LEGAL&NPAY, which carries its own impact (in the form of PW and NW) upon the degree of belief in the goal hypothesis.

After interactions had been inserted, to accomplish further refinement the experts operated the system in "full trace" mode (so called). Although this mode was slow in operation, even tedious, it allowed the experts to observe the effect upon the hypothesis which might be caused by alternative possible responses of a user, thus stimulating them to suggest changes in the values of the parameters, PW and NW. The goal sought by these changes was to calibrate AUDITOR to the extent that a probability of at least 90%, (DB 9.5) should be reported by AUDITOR at the time of the expert's own report of a subjective feeling of "satisfaction" in his judgment, despite the difficulties with this procedure which might be predicted by a review of recent research, c.f. Lichtenstein, Fischhoff, and Phillips (1977), Crosby (1981), and others, summarized both in Libby (1981) and in Ashton (1982). Although consensus was found among the experts in the direction of the changes which they suggested, the model proved somewhat insensitive to the exact values to which the changes were made, a result reminiscent of the work of Dawes (1975) and of Einhorn, Kleinmuntz, and Kleinmuntz (1979), a matter which will be explored in a subsequent paper. The complete set of rules and parameters utilized in the system may be examined in Appendix [C].

VALIDATION

The results of AUDITOR's operations were compared against the judgment of practicing auditors utilizing client audit data contained in work papers. Two validation procedures were employed: "Open-book" and "Blind."

OPEN BOOK: The validation procedure referred to as "Open-book" was performed by an audit manager in a different office of the same CPA firm which participated in the building of the system. This auditor, serving as validator, selected work papers covering completed audits of two commercial clients with which he personally was not involved. From the work papers of each of the two clients the validator chose five individually large, delinquent accounts. The criteria which he used for his selection was not communicated to the researchers. Each set of work papers to which he referred contained the decisions made by the auditors during the audit about their assessment of the collectibility of these five delinquent accounts (and others). Since the validator was unfamiliar with the two clients which he had selected, the work papers presumably also contained all the information from which he determined his responses to AUDITOR's questions. Since to the validator it was obvious which set of judgments appeared in the work papers and which "judgments" were produced by AUDITOR, this procedure is referred to as an open-book validation.

After a demonstration of operation of the system by the researchers, the validator invoked the AUDITOR program and responded to its questioning by entering information which he gathered from the work papers. That is, he responded to its inquiries by entering a Certainty Value on a scale from -5 to 5 in answer to each question presented to him following the prefix, "How certain are you...?"

In the case of each delinquent account, the validator himself decided when to exit from the system, thus concluding one session with AUDITOR. Presumably, that occurred either when he believed that he had provided to AUDITOR enough data by which he personally could have made a decision concerning collectibility of each account, or when all data available in the work papers had been entered into the system.

At the conclusion of each session the validator relayed to the researcher the judgment of the CPAs who had performed the audit. He also examined AUDITOR's report of degrees of belief and probabilities which had been produced as a consequence of the interactive question and answer session. On the basis of his own criteria he labeled each result a "hit" (appropriate result) or a "miss." (In every case he commented that his own judgment regarding the write-off of the delinquent account agreed with that reported in the work papers.) The results are summarized in Table 1.

RESULTS: In each case but one, when the validator chose to terminate questioning and call for the report by the system, AUDITOR reported a degree of belief which was the equivalent of a probability of at least .865 in favor of the same decision as that called for by the original audit team--that is, either in favor of, or against, the need for an allowance for the delinquent account under scrutiny. The validator and the researchers considered these results to be successful.

BLIND VALIDATION: A second validation procedure was carried out to conform to the suggestions made by the mathematician A. M. Turing (1950).

TABLE 1

SUMMARY OF RESULTS OF FIRST VALIDATION

Debtor	Description	(Probabilities Reported by)		No. of rules invoked by AUDITOR and its user	CPA Firm's Decision as Reported in its Workpapers
		AUDITOR System For RESERVE	Against RESERVE		
Client #1 (Steel Mill)					
1	\$7,000 delinquent; \$2,000 sub- sequently collected		.99*	6	No Reserve
2	\$78,000 delinquent out of \$457,000 total owing	.96		5	\$37,000 Reserve
3	\$51,000 total account, \$15,000 delinquent; claim for damaged goods	.997		5	Reserve \$15,000
4	No detail recorded by the researcher	.999		2	Reserve
5	No detail recorded by the researcher		.865	11	No Reserve
Client #2 (Heavy Equipment Manufacturer)					
1	All current except one large delin- quent charge	.988		7	Reserve
2	All collected subsequent to balance sheet date		.999	2	No Reserve
3	Debtor is out of business	.93		5	Reserve
4	\$48,000 delinquent		.865	12	No Reserve
5	Debtor is bankrupt	.93		5	Reserve

*Decimals are probabilities converted from the "degrees of belief" scale used in AUDITOR. 1.00 would represent certainty. Thus, .99 Against means the AUDITOR system advises the need for providing a reserve for this particular account; the evidence, as furnished by the user of the system, stacks up 99% in favor of this result, that is, that no reserve is necessary. The column headed "CPA Firm's Decision" reports the judgment reached by the auditors.

This procedure is considered to be a more stringent test of the validity of an expert system and has become somewhat of a standard in the field (c.f. Yu, 1979). Turing, when confronted with the issue of whether computers and their programs could be said to think, proposed instead that the testable question is whether an observer who is himself ignorant--that is, blind--as to the identity of the source could distinguish between the output of the machine and that of a human.

Blind validation of AUDITOR utilized two human auditors, one who served as user of the system, the other as validator. Both auditors were managers with an international CPA firm not previously involved in the AUDITOR project. The auditor who performed as user selected an audit client with which he was familiar. His choice was a NYSE listed company which manufactures and markets through various distribution channels consumer products such as cigarettes and candies. His audit team, in their visit to the client's offices at an interim date, had identified by computerized selection eleven invoices so large and delinquent as to justify the auditors' individualized attention. The criteria for this selection was not communicated to the researchers.

The user was familiarized with the AUDITOR system and then instructed to respond to its questions using the data contained in the work papers. Since he himself had been a member of the audit team it is not possible to determine if his answers were based solely upon the contents of the work papers or perhaps also relied upon his memory for information not in the papers. He answered the system's questions, responding on the CV scale from -5 to 5. As in the first validation, the user himself decided when to terminate each session. Concurrently, he reported relevant data from the work papers to the researcher who recorded it, such as age, dollar balance, confirmation results, nature of customer's business, etc. This data--recorded on Fact Sheets--became part of the raw data presented to the validator from which he made his judgment. The user also reported the audit teams' judgment, just as it was recorded in the work papers. This information was recorded by the researchers on a Comparison Worksheet, on which was also entered the "judgment" of the AUDITOR system.

Subsequently, the audit manager serving as validator was presented with the Comparison Worksheets which contained the judgments from the two sources and the Fact Sheets containing the data from which they made their judgments. (Of course, the sources of each judgment were disguised when presented to the validator; that is, he was "blind" as to the source--auditor or AUDITOR--of each judgment.) The task which he undertook was to study the same evidence on the Fact Sheets which had been available to the Fact Sheets which had been available to the two experts, auditor and AUDITOR, to form his own independent and presumably expert conclusion, and then to accept or reject each judgment from each source on the basis of his opinion of the expertise which each had demonstrated. That is, he accepted or rejected each of twenty-two recommendations which were disguised as to source.

RESULTS: The validator accepted all of the audit team's judgments and all but one of AUDITOR's. These results are summarized in Tables 2 and 3. In sum, over the two validation procedures, open-book and blind, the AUDITOR system scored nineteen hits out of twenty-one cases.

ANALYSIS

TAXONOMY

In AUDITOR a high score on one cue can offset a low score on a different cue. Thus, AUDITOR can be called a compensatory, rather than non-compensatory model. For example, credit to the delinquent customer may not have been stopped by the client (CREDITSTOP), and the customer may have demonstrated a good record in the past of paying his account (GOODRECORD). However, both of these cues, favorable to an expectation of collectibility of the account, may be offset--compensated for--by evidence that the customer is in bankruptcy proceedings (BANKRUPTCY) and the client's lawyer advises that recovery is unlikely (LAWYER).

AUDITOR incorporates probabilistic aspects of decision making, as do other successful expert systems, recognizing that no evidence is perfectly diagnostic of the condition which the expert must evaluate and that real-world judgments are made on the basis of information which must be probabilistically evaluated. The query format, "How certain are you...?" reminds the system's user to perform a subjective assessment of the reliability of the evidence.

In contrast with Lens Model studies, and indeed with most studies of auditors' judgments except those utilizing Verbal Protocol Analysis, AUDITOR processes the independent variables with which it operates, in the form of user's CV applied to the PW or NW of a cue, in a sequential fashion. That is, AUDITOR continuously revises and expands the system as its queries are successively put to and answered by the user. In sum, AUDITOR incorporates subjective assessment of the evidence through what is basically a Bayesian revision process operating on cues which can compensate for each other. Therefore, the AUDITOR system can be called a subjective, Bayesian, compensatory expert system, which functions as a sequential model of auditors' judgment.

COMPARISON WITH OTHER STUDIES

Two previous studies have investigated auditors' perceptions of what was called source reliability (Joyce and Biddle, 1980) or source credibility (Bamber, 1980, quoted in Libby, 1981). Joyce and Biddle explored whether auditors weighted differently the reliability of information depending upon its origin with the client's credit manager or with an outside credit reporting agency. Joyce and Biddle concluded that unless both sources were called to the auditors' attention (as was done in a within-subjects experiment) they did not differently weight the sources. In the AUDITOR project no rule incorporating outside credit reports was suggested by the experts, and apparently none was referenced

TABLE 2

CONVERSION OF DEGREE AND PROBABILITY INTO VERBAL JUDGMENTS

<u>Case #</u>	<u>AUDITOR was Concealed as Expert No.</u>	<u>AUDITOR's Reported Degree of Belief and Equivalent Probability</u>		<u>AUDITOR's Result Was Reported to Validator as</u>
		<u>DB</u>	<u>%</u>	
3-1	II	-15.5	3.2	"OK, no res. nor adj."
3-2	I	-19.5	1.3	"No reserve nor adj. needed"
3-3	II	15	97.0	"Yes, reserve or adj. req'd"
3-4	I	-30	.01	"No"
3-5	II	-17	2.0	"No"
3-6	II	26	99.0	"Yes"
3-7	II	2	61.0*	"At most a partial res. or adj. req'd"
3-8	I	-30	.01	"No"
3-9	II	-30	.01	"No"
3-10	I	-30	.01	"No"
3-11	I	33	99.9	"Yes"

*There were no clear guides for interpretation of the implications of AUDITOR's result at this level of likelihood (61% probability of need for an allowance). Since the first validation procedure could be interpreted as suggesting that partial allowances by auditors might be associated with the AUDITOR's results at this level of probability, the researcher chose to state the expert's opinion as "At most a partial res. (reserve) or adj. (adjustment) (is) req'd (required)".

TABLE 3

SUMMARY RESULTS BLIND VALIDATION

Case #3-	Brief Description	-----Judgments-----		Validator Called
		AUDITORS	Auditor	
1	\$19,292 over 90 da. active gov't account, not confirmed.	No reserve	No res.	<u>Accept</u> A
2	\$17,553 over 90 da. active gov't account, not confirmed.	No res.	No res.	A A
3	\$16,346, 30-60 da., 100% credit memo not posted. Not confirmed.	Res. or adjust	Reserve	A A
4	\$35,909, 30-60 da., not confirmed. Workout agreement in file.	No res.	No res.	A A
5	\$59,570, 30-60 da., not confirmed. Bankrupt but letter of credit in file.	No res.	No res.	A A
6	\$13,563, 30-60 da., not confirmed. Customer request to return goods, C.M. not posted.	Res. or adjust	Res. or adjust	A A
7	\$11,628, over 90 da., confirmed newer items no excep- tion. Creditmanager calls excellent account, but dispute over price.	At most, partial reserve or adj.	No res.	Reject A
8	\$12,498, 30-60 da., not confirmed, later collected.	No res.	No res.	A A
9	\$14,655, over 90 da. not confirmed, later collected.	No res.	No res.	A A
10	\$22,024, over 90 da. not confirmed, later collected.	No res.	No res.	A A
11	\$36,763 delinq. and total. Letter in file: "Financial difficulty, no pay next six months."	Yes, res.	Yes, res.	A A

in the work papers--which covered large, commercial clients--consulted by the auditors during validation of the system. Therefore, no direct comparisons with Joyce and Biddle are possible.

However, several other outside data sources were included in rules suggested by the experts involved in building AUDITOR. Among these are PROBLEMS, relating to the results from confirmation requests, LAWYER, concerning the opinion of outside legal counsel, NOTPAY, which reports the debtor's stated intent not to pay (when that is known), and NONCONTACT, which is a cue activated by the client's and the auditors' inability to contact the debtor. These cues carry four of the five heaviest, basic weightings (PW or NW) in AUDITOR. It seems reasonable to believe that these relative diagnostic weightings carry within them a factor which is dependent upon the experts' belief in the reliability of the source of the information. Thus, these heavier relative weightings evidently reflect the experts' belief in the higher reliability of these outside data sources.

Additionally, the low diagnostic weight accorded to the rule CREDITMGR, which reports the opinion of the client's credit manager, seems to indicate that such a source, although frequently resorted to, carries low reliability. Bamber's 1980 study attempted to determine if audit managers differentially weighted the results of work done by audit seniors who were described in the experiment as being of differing reliability. No comparison with the AUDITOR project is possible, since, as might be expected, no cues suggested by the experts focused on the competence of the audit staff.

RELATIVE IMPORTANCE OF THE RULES

Cues frequently used which also have heavy diagnostic weight are, by implication, the major diagnostic tools of the experts. Table 4 lists the cues in decreasing order of their importance in use as diagnostic indicators during the validation trials. This table has been prepared, first, by totaling for each cue the largest of its Positive Weight or Negative Weight together with the weights of the interactions involving that cue (all in unsigned terms). Then, this combined weight has been multiplied by the number of uses occurring during the validation trials. (Cues with zero usage during validation--implying that no evidence concerning the cue was present in the working papers--are excluded from this table.) The result provides a measure of the extent to which the system utilizes the different cues.

Some comments are appropriate concerning the frequency of usage. The figure for number of uses is a count of the number of times during validation in which a non-zero Certainty Value constituted the response by the system's user. Since twenty-one delinquent accounts were examined during the validation runs, that number is the maximum which could appear as number of uses. The cue called COLLECTED serves less as a diagnostic tool for the experts than as a means within the AUDITOR system to identify which accounts continue to be doubtful, following

TABLE 4

CUES, THEIR WEIGHTS AND NUMBER OF USES, AS DIAGNOSTIC INDICATORS

Cue	Largest Unsigned PW or NW	Combined Weight W/Inter- Actions	No. of Non-Zero Uses	Combined Weight X Uses
COLLECTED (Account is no longer delinquent by audit completion date)	30	30	20	600
ECONOMIC (Economic factors hamper ability to pay)	3	19	9	171
BANKRUPT (Debtor has entered bankruptcy proceedings)	3	19	8	152
CORRESPOND (Recent data in credit file supports collectibility)	4	11	13	143
ACTIVE (Debtor continues as active customer)	3	15	7	105
WORKOUT (Recent collections are proceeding satisfactorily)	3	6	13	78
NOTPAY (Debtor expresses his intent not to pay)	5	14	5	70
PROBLEMS (Confirmations reveals serious problems)	8	11	5	55
CREDITMGR (Client representative expresses strong belief in collectibility)	2	5	9	45
LAWYER (Lawyer expects poor prospects of recovery)	5	21	2	42
LEGAL (Debtor has strong counterclaim)	3	9	4	36
WRITEOFF (Writeoff of this account represents a material adjustment)	2	11	2	22
COLAGENCY (Collection has been assigned to agency or lawyer)	2	7	3	21
GOODRECORD (Debtor has good record of paying in past)	2	4	5	20
NOPAYEVER (Debtor has made no payments on any invoice)	2	10	2	20
NONCONTACT (Confirmation request was returned undeliverable)	4.5	9.5	1	9.5

their initial identification and listing as delinquent, perhaps during an early stage of the audit.

Rules with the highest usage reflect aspects of the auditors' apparent decision processes and usage of evidence. WORKOUT reflects the significance to the auditor of his knowledge of recent cash collections from the delinquent debtor. The phrase, "proceeding satisfactorily," in effect asks the auditor to evaluate the recency, regularity, and adequacy of the collections in comparison to what he judges is some acceptable standard. Because of the crucial nature of the sales and collection cycle, the inspection and testing of subsidiary accounts receivable ledgers and cash receipts records normally would be part of every audit. Thus, the information required in WORKOUT ordinarily would be available in every audit.¹⁰

The frequent usage of CORRESPOND similarly reflects typical audit practice--that of examining whatever correspondence is available concerning a delinquent debtor. To some extent, also, this cue constitutes a catch-all. The great variety of different data which might be included in correspondence files by different clients and for different debtors discouraged the researchers from attempting to incorporate in AUDITOR's rule base each separate cue which might be discovered. For example, one cue might indicate that a delinquent debtor has placed with the client a bank's letter of credit, which effectively removed any doubt about ultimate recovery of the account. A similar but different cue to be found in another debtor's files might refer to a parent company's guarantee of indebtedness. To avoid a proliferation of such rules in this first attempt at building an expert system of auditor's judgments, CORRESPOND was formulated to allow the user to sum up in one response the data which he gathered from the client's correspondence files. Such a response constitutes in effect a "mini-judgment" by the user, which more experience with the system may suggest should be divided into several rules. ECONOMIC is also a rule of rather frequent usage, perhaps as a result of concern about the economic conditions which prevailed when this work was done in (1981 and 1982). Usage of CREDITMGR reflects the common audit practice of discussing delinquent accounts with a representative of the client.

On the other hand, several rules carry a relatively heavy impact but evidently are infrequently applied. Auditors report they seldom communicate with an attorney regarding a delinquent account, but would give strong weight to a pessimistic prediction which he might make (LAWYER). The creditor's stated intent not to pay could represent an important cue but is rarely available (NOTPAY). The only instances reported to the researchers of the presence of this cue occurred in the case of debtors who were disputing the propriety of the charges in their accounts, claiming errors in pricing. Similarly, if neither the auditor nor his client were able to communicate with the debtor that would represent an important cue (NONCONTACT). No such situation arose between the large commercial and government entities involved serving as test cases in the AUDITOR study.

FURTHER ANALYSIS

Three other aspects of the system and its usage are worthy of comment. Michie (1980) and others have suggested that an experienced practitioner's expertise consists of his ability to recognize clusterings of cues--that is, patterns--which he has encountered previously and found to be particularly diagnostic. Clusterings of cues may function as mini-hypotheses, allowing the expert to focus his evidence-gathering in an economical fashion. To the researchers, the most obvious of the patterns apparently utilized by auditors in the valuation of accounts receivable seems to be organized around a distinction between "slow-pay" versus "no-pay" as diagnostic categories for each delinquent account. Thus, a slow-pay customer, properly diagnosed, does not require a provision for loss. However, this categorization may hide several patterns of less global impact, which may in reality be the organizers for a practitioner's judgment, such as "the economic-problem" pattern or "the legal-problem" pattern. This matter will be explored in a subsequent paper.

The second matter concerns strength of belief and the meaning of AUDITOR's report. AUDITOR reports Degrees of Belief. This constitutes its expert judgment. The user of the system must decide for himself the significance of the report, for example whether 10 Degrees (.91 probability) mandates that an allowance be provided. The researchers attempted to calibrate the system during its refinement stage so that AUDITOR would produce a report of about 10 Degrees (positive or negative), coincident with the time a human expert auditor reported that he had become "satisfied" with the data and had made his decision. However, many studies reflect unfavorably on humankind's ability to properly deal with data presented in a probabilistic format, c.f. Joyce and Biddle (1981a). Thus, AUDITOR's report of Degree of Belief can best be viewed as an indication of relative strength of belief. No claim is intended either that AUDITOR's report reflects objective probabilities nor that the researchers believe human experts process probabilities similarly in the construction of their judgments. However, while keeping this disclaimer well in mind, the researchers believe that by the time the users terminated their validation sessions AUDITOR can fairly be said to have reached an unambiguous result, at least in terms of the refinement criteria, in the great majority of cases. For example, for the third client, which comprised the blind validation trial, in only one instance of eleven delinquent debtors was the session terminated when AUDITOR was reporting less than four percentage points from a value which would represent certainty (0% or 100%) Table 2. In the ten open-book trials, eight reported within +/- seven points, results which are within the criteria (Table 1). Thus, the researchers believe this first complete version of AUDITOR can be said to diagnose the collectibility of delinquent accounts in an unambiguous fashion, which bodes well for the development of other expert systems for use in audit situations.

The final matter for discussion concerns the sequence of inquiry followed by the system. AUDITOR asks next that particular question

which potentially has the greatest impact upon the probability of the hypothesis. Human auditors do not organize their own evidence-gathering in such an efficient fashion, nor do they plan the audit in order to investigate every source of highly reliable information. For example, it is not a standard audit procedure to request confirmations from every large delinquent debtor, even though responses might reveal evidence of high diagnosticity through application of such rules as PROBLEMS, NORESPONSE, NONCONTACT, LEGAL, and NOTPAY. Neither is it common practice uniformly to inquire of the client's attorney or collection agency concerning a delinquent debtor, although evidence of high diagnosticity might result (LAWYER, BANKRUPT, LEGAL, and various interactions). Explanation for this apparent gaffe is easily found. Auditors are cost conscious, AUDITOR is not. This expert system is designed to concentrate on the diagnosticity of the evidence while ignoring its cost. Human expert auditors, at least those who work for profit-oriented firms, are expected to be time and cost conscious.

Additionally, of course, an eclecticism in auditors' pursuit of evidence is motivated by factors such as the unpredictable availability of clients' records and employees, conflicting schedules necessitated by a desire to service several clients, and a need to accomplish several audit objectives in a limited time. Also, since both client and auditor seek more from the audit than a satisfactory valuation of the allowance for bad debts, other tasks may intervene.

In only one case during validation did sequencing of inquiry cause difficulty. In that case, #1-2, the user serving as validator terminated the validation session immediately prior to a question which had great bearing on that particular debtor and which would have put AUDITOR more in agreement with the judgment of the human auditor, that is, at a higher Degree and probability level. In all other instances, the sequence of greatest-impact-first, with termination controlled by the user, appeared to be a satisfactory method of handling the evidence contained in working papers without provoking a complaint by the user of the system. Thus, the sequencing in AUDITOR seems to be at least satisfactory to its users in the majority of cases. Whether it is reasonable to expect auditors to follow a more efficient search pattern is beyond the scope of this paper.

SUMMARY

Expert systems such as AUDITOR can constitute a model of auditors' judgment, answering the researchable and interesting question said to be at the heart of all investigation of diagnostic judgment in the face of uncertainty: what cues do the judges use, what are their weights, and how are they assembled into a judgment model. Thus, subject to the influence of the researchers' intervention, AUDITOR demonstrates the feasibility of a new approach to the descriptive study of auditors' judgments.

LIMITATIONS

All of the auditors who participated in the initial system-building were members of one office of one large public accounting firm. To a great extent all had been exposed to similar professional training. Additionally, they perform under policies which might perhaps be unique to one firm. However, auditors from other firms were involved in the refinement and validation stages of the project--which lends confidence in the universality of the system as a model of auditors' judgment and reduces the chances of parochialism.

Further refinement of the system may produce better results. Rules might be refined to reduce them more nearly to the elemental cues, reducing the need for the mini-judgments presently contained in such rules as CORRESPOND and LEGAL. Such a revised rule base might be more reliably applied by an unexperienced auditor--one who had not developed competence in the mini-judgments which a few of the rules call for. On the other hand, such changes increase the size of the rule base and lengthen the list of questions to which a user is exposed--factors which tend to decrease the convenience of working with the system. However, work papers used by the auditors who were involved in various stages of this work evidently contained very little relevant data beyond that called for in the rules of the system, with the exception of the age of the delinquent balance which was under scrutiny. Some indication during validation was gained that the absolute age of the account may have to some auditors a significance not reflected in the system. Therefore, a rule relating to age of the delinquent balance might be useful in the system. However, for the large, successful, client companies whose audit work papers were referred to in various stages of this work it seemed likely that no delinquent account was allowed to gain more than a few months' age before some finality was forced in its disposition. (These can be presumed to be clients with relatively good internal controls over sales and collections.) However, there is the possibility that alternative systems of rules and weights may produce equally successful systems which demonstrate audit judgment.

The influence and beliefs of the observer--the researchers--cannot be ruled out in project of this kind. Also, the subjects' desires to please and accommodate the researchers may work to the detriment of effort to create and particularly to validate an effective expert system.

Criticism can be directed against the system for its use of Bayes' revision. It is commonly believed that auditors are no more Bayesian in their processing of evidence than are other humans (Libby, 1981). However, the point is worth emphasizing that the use of Bayesian revision in computerized expert systems operating in fields other than auditing produce results similar enough to those produced by human experts as to satisfy these experts themselves.

Further refinement of the system may be called for, particularly, regarding instances in which practicing auditors conclude that partial rather than 100% allowances are proper to provide against a delinquent

account. In the Open-Book Validation procedure, auditors called for a partial allowance against an account for which AUDITOR's processing produced (upon replay by the researcher) a result of approximately 70%. Further experience with the system may serve to clarify the set of circumstances under which human experts call for partial allowance and the probabilities which AUDITOR might be expected to report in the same circumstances.

On the other hand, the judgments of expert auditors may be sufficiently idiosyncratic and inconsistent that attempts to refine the system any further will be fruitless. Changes to the system will be carried out only after an analysis of the objective to be sought. For example, if enhanced convenience of the system for use by practicing auditors were to be an objective of modification, a desirable change might be to engineer a revision in the order in which certain questions are asked. Certain evidence, while of lower diagnostic impact, seems always to be available, and might be asked about early in a consultation to allow the user promptly to contribute his knowledge early in questioning. On the other hand, already existing capabilities within AL/X permit the user to volunteer evidence prior to questioning. For example, if no request for confirmation was sent and no information was received from any outside source, the system could be modified to accept this data even before questioning began. These capabilities have not been exercised yet in AUDITOR.

FOOTNOTES

¹Named second in frequency was the judgment process leading to the budget of audit time, the study of which might have allowed useful comparisons with previous research (c.f. Joyce, 1976). However, the widely differing descriptions of this process which were given by the experts discouraged the researchers from attempting to extract their expertise. The judgment process involved in determining the threshold for "reporting materiality" was similarly considered then rejected for study because the aspects of negotiation evidently resorted to between client and auditor have not been amenable to expert system technology.

²AL/X (Advice Language / X) was developed for Intelligent Terminals Ltd. by John Reiter, Steve Barth, and Andy Paterson in association with the University of Edinburgh and was supported by BP Petroleum Development Ltd. It is a Pascal system based upon the Prospector consultant system for mineral exploration developed at SRI International by Richard Duda, Peter Hart, and others, see Duda et al. (1979).

³The capabilities of AL/X allow formation of queries in a form calling for a "Yes" or "No" answer. This capability was installed after the work on AUDITOR had begun. Because of the apparent ease with which users of the system had already adapted their responses to the CV scale, this added capability was not used in AUDITOR.

⁴RESERVE does not make provision for situations in which the auditor concludes the amount to be provided for potential loss of the account should be more than zero but less than 100% of the delinquent balance. One such case was encountered among the twenty-one delinquent accounts reviewed during validation of the system, and the validator ruled that AUDITOR did not perform satisfactorily.

⁵Four separate sets of cues were accumulated from these experts. The nature of differences between these cue sets will be explored in a later paper. The final AUDITOR system was an aggregation of the cues provided by the four experts.

⁶Rules related by AND linkages are processed according to the "fuzzy logic" rules of L. Zadeh (1979): $P(A1 \text{ AND } A2 \text{ AND } \dots \text{ AND } A_n) = \text{minimum}[P(A1), P(A2), \dots, P(A_n)]$, where P is probability and the A_n are evidential cues.

⁷It was also necessary to "prime" the system by providing the Bayesian portions of it with initial (prior) degrees of belief. In the absence of any definitive guidance in the literature or from the experts all such initial values were set to reflect 0.0 degrees of belief, which is a probability of 0.5. At the beginning of any session with AUDITOR these values can be changed if desired.

⁸In both open-book and blind validation procedures the assumption was made that the work papers contained both the auditors' judgments and the support therefore, in conformity with Generally Accepted Auditing Standards, AU 338.

⁹Since AUDITOR reports its results in degrees of belief and in probabilities, which would be a novel way for a subjective audit judgment to be recorded in work papers, it was necessary for the researchers to interpret and in effect to translate AUDITOR's report before its presentation to the validator in order to maintain anonymity of the source of each judgment appearing on the Comparison Worksheet. This interpretation was easily made in all but one case. In every case but one, AUDITOR's report produced a DB of at least 15, positive or negative, that is, probability of at least .97 either in favor of or against the hypothesis which called for an allowance. These results appeared to the researchers to be unequivocal and were reported on the Comparison Worksheet as that of an expert auditor who called for "No reserve nor adjustment needed," (in the cases which reported negative DBs) or "Yes, reserve or adjustment needed" (in cases which reported positive DBs). (In keeping with common practice the term "reserve" was freely used as a synonym for allowance for bad debts among the researchers and the practitioners.) However, one case, that of number 7, produced a DB of 2 (probability .61)--an equivocal result having neither precedent nor clear interpretation within the researchers' experience with the AUDITOR system. The researchers reported this on the Comparison Worksheet as, "At most a partial reserve or adjustment is required." On the same case, the audit team had found no need for an allowance; the validator scored this judgment at a hit. The validator initially scored AUDITOR's result as a miss and that is the way it was recorded by the researchers. Later, however, when reviewing his work, he referred to AUDITOR's result on number 7 as a "close call." The issue of partial allowance will be studied further in later applications of AUDITOR.

¹⁰This cue contributes to the researchers' impression that the expert auditors' basic diagnostic task consists of an attempt to classify each delinquent account as "slow-pay" or "no-pay," that is, these categories constitute competing hypotheses for classification of each delinquent account.

¹¹Technical data for this appendix also came from "AL/X USER MANUAL," Andy Paterson, Intelligent Terminals LTD. Oxford, England, 1981.

APPENDIX A: Description of AL/X and AUDITOR¹¹

The structure of an expert system, in general, and AL/X, in particular, parallels the structure of human decision making. Data is gathered, analyzed, evaluated with respect to some criteria and used, and, if more data is required, what additional data to gather is determined. Figure A-1 shows how the specific step in the operation of AL/X map to the generic decision making phases.

The two key operational areas are revision and selection. Revision refers to the process of updating the Degree of Belief associated with the spaces, that is, the rules and hypotheses. Figure A-2 depicts this revision as a two stage process. At all times the hypothesis, which is the goal of the system's inquiry process, has associated with it a Degree of Belief (DB) or simply degree. Degrees of Belief measure strength of belief on a scale having a range from -100 to +100 but, AUDITOR only uses a range of -30 to +30. The initial ("prior") DB value within AUDITOR of the hypothesis, RESERVE, is set to 0.0 but can be easily set to any initial value. The DB scale is derived from probabilities as:

$$\text{Degree of Hypothesis(H)} = 10 \log_{10}(\text{Probability(H)/(1-Probability(H))})$$

Similarly, every space carries a DB which reflects its prior strength of belief. The user's response (CV) to each question initiates stage 1; revision immediately of the DB for that space (unless the response is 0). The value of CV revises the degree in the following manner. A user's response of +5 assigns to that space the maximum DB permitted by the systems, i.e., +100. Similarly, an answer of -5 results in a degree of -100. An answer of zero leaves the degree unchanged at its prior value. An answer other than +5, -5 or 0 results in interpolations between these three points. A linear interpolation is performed on probability values. A positive CV causes interpolation between the prior probability and 1. A negative CV value results in interpolation between zero and the prior probability. The resulting probability value is then converted back to a degree value. The effects of the user's answer are immediately propagated through the inference net to update the DB of all spaces for which the selected question is evidence, including the hypotheses.

The second stage of the two-stage process of revision involves calculations of the incremental weight which will be added to the present degree of the hypothesis (and to the present degree of any other spaces which, in a particular system, might be consequences of the immediate question to which the user is responding). First, we will speak about revisions involving spaces connected via IF:THEN inferences. A process of interpolation is again used employing the PW or the NW of the space, together with the degree calculated from the user's CV response. If the evidence is true then a positive weight (PW) of one will increase the degree of belief of the hypothesis by one. Similarly if the evidence

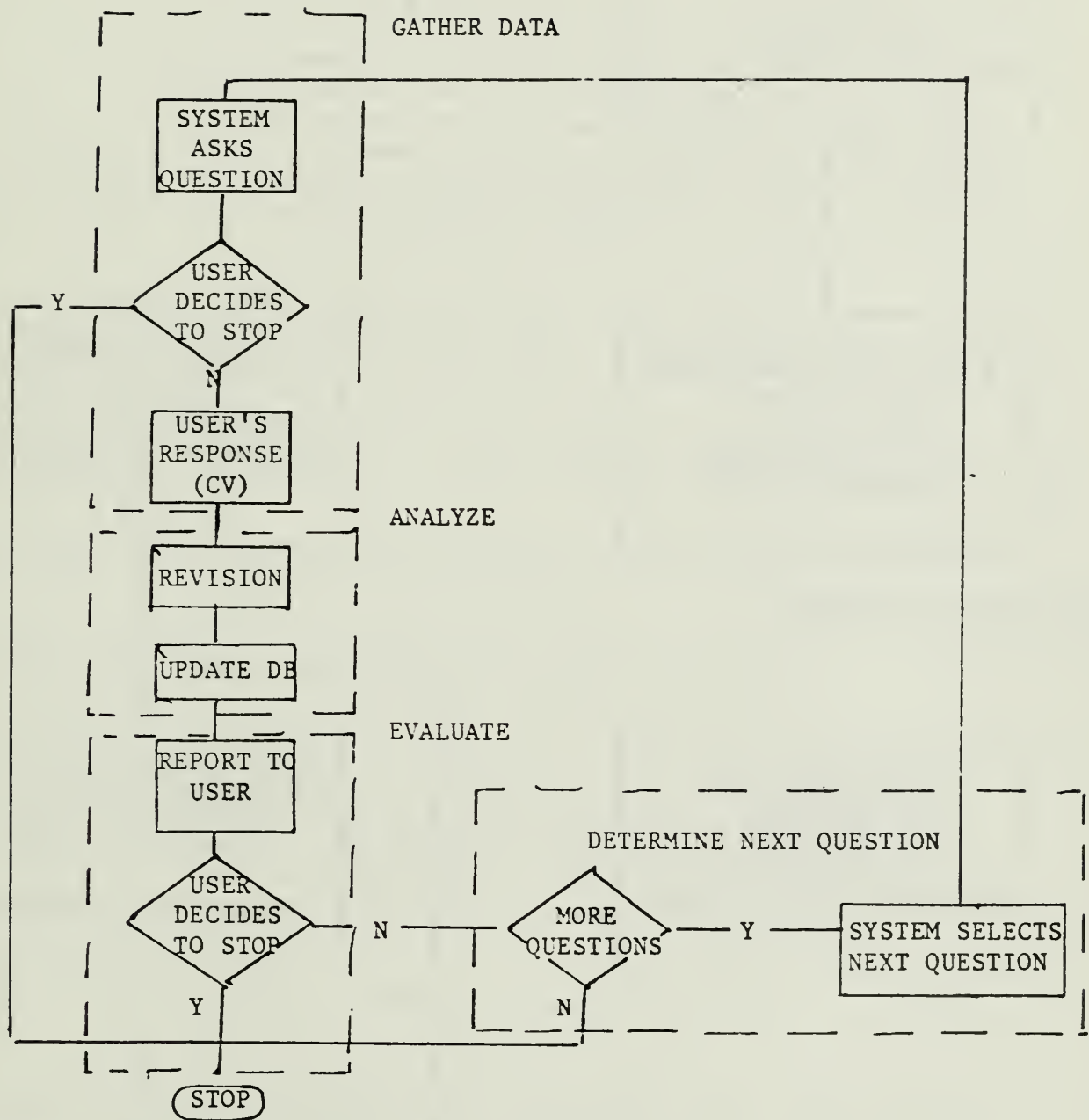


Figure A-1 Overview of Expert Systems

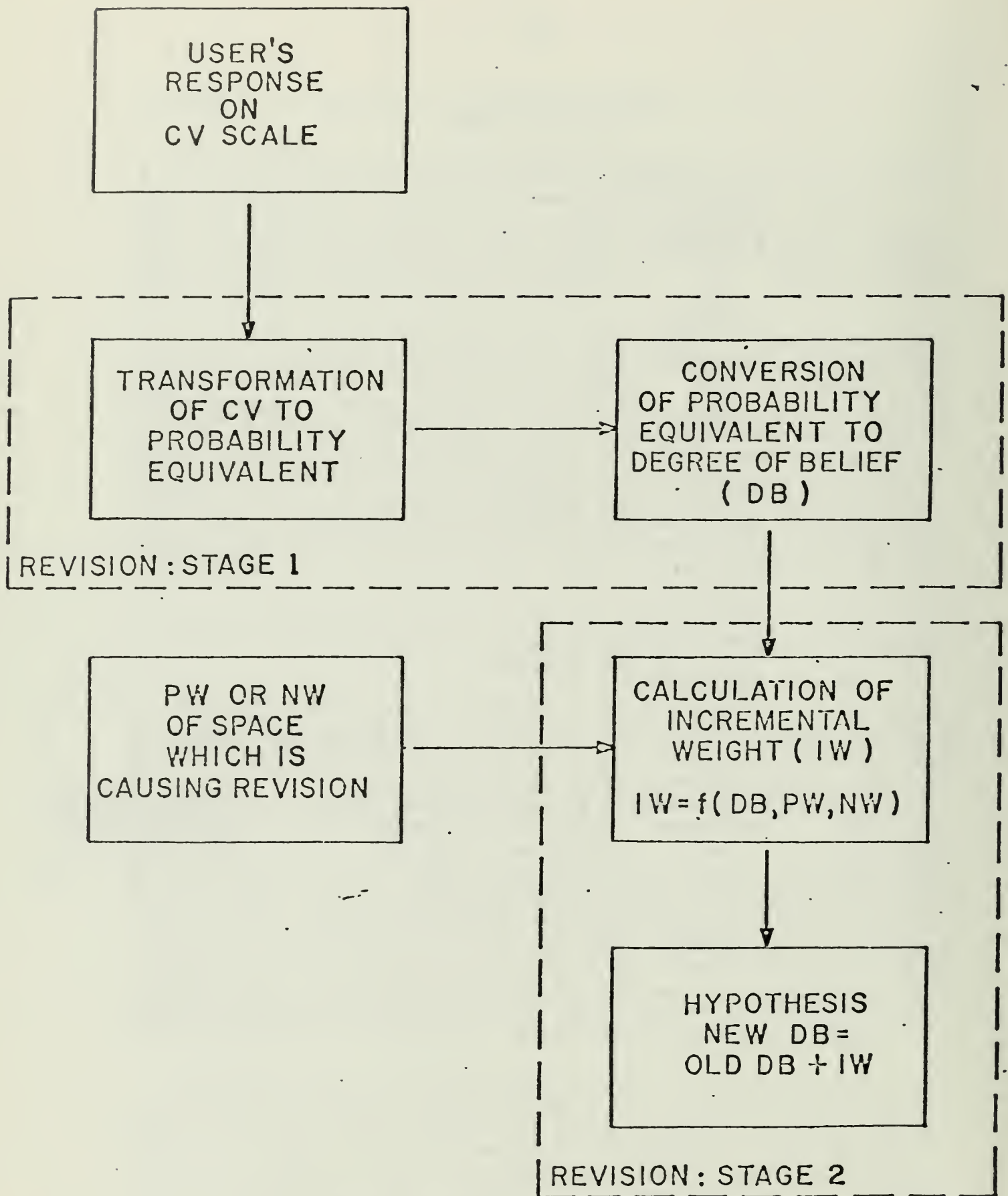


Figure A-2 Overview of the Process of Revision

is false and the negative weight (NW) of one then the degree of the hypothesis will decrease by one. The greater the positive weight then the more strongly does the presence of the evidence imply that the hypothesis is true. For negative weights, the corresponding implication is that the absence of the evidence implies that the hypothesis is false. In the situation where the presence of evidence implies that a hypothesis is false rather than true it is perfectly legitimate to have a negative, positive weight and a positive, negative weight (i.e., $PW = -4$ and $NW = +5$).

If the degree of the space, as a result of the user's response is greater than the prior degree of that space, then the incremental weight to be added to the degree of the hypothesis is calculated as

$$\text{Incremental weight} = PW / (PW * \min(PW, \text{current} - \text{prior}))$$

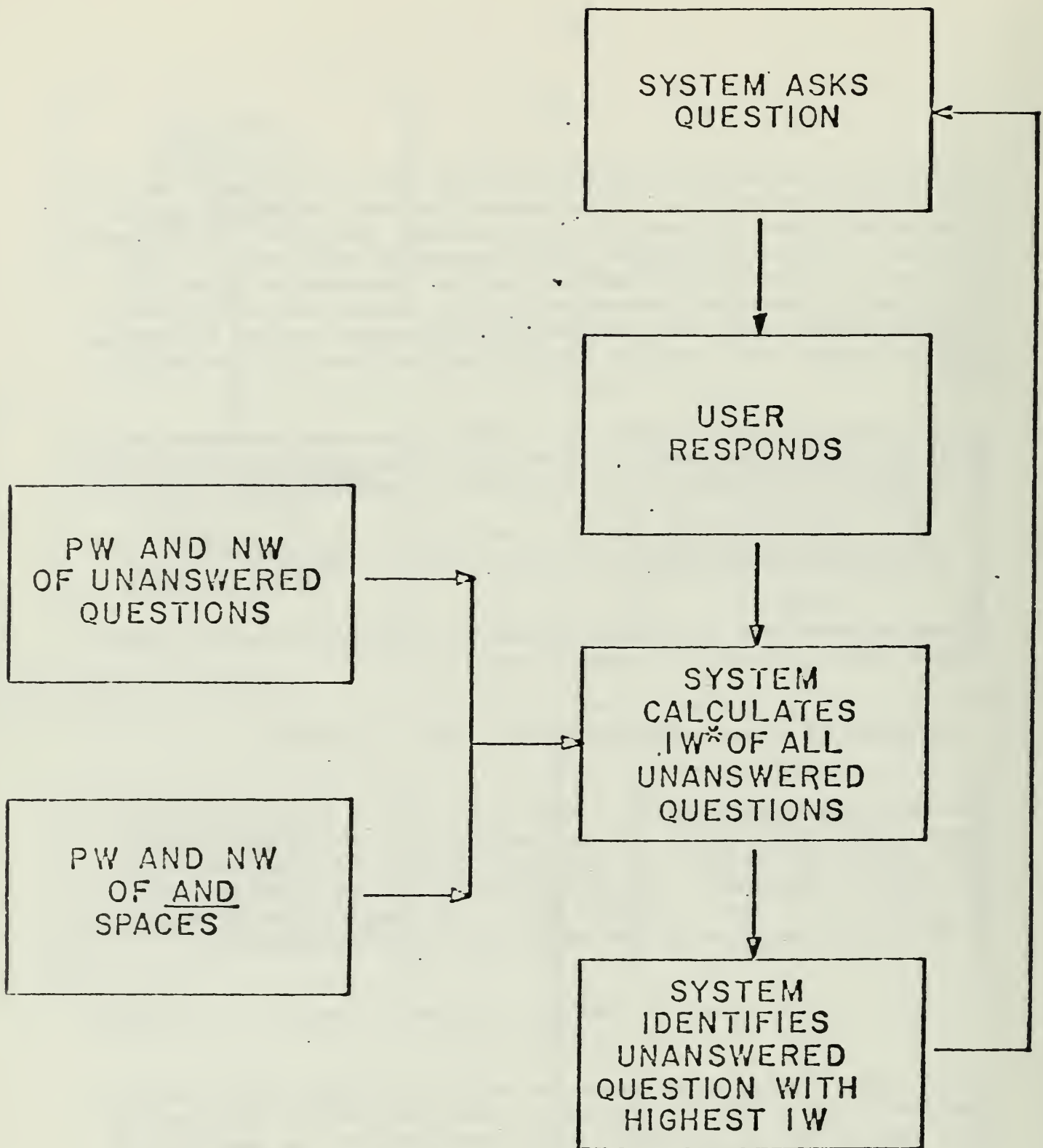
If the degree of the space as calculated is less than the prior degree of that space, then the incremental weight is:

$$\text{Incremental weight} = NW / (NW * \min(NW, \text{prior} - \text{current})).$$

AND spaces are composed of the connection by AND of two or more IF:THEN spaces. Each AND space has its own degree of belief. Also, each AND space has its own PW and NW. The weight which an AND space has on the hypothesis is dependent on the degree of that AND space and its PW and NW, just as if it were an IF:THEN space. The degree of an AND space is a function of the degree of the components and is calculated as

$$\text{degree}(A1 \text{ AND } A2 \text{ AND } \dots \text{ AND } A_n) = \min(DB(A1), DB(A2), \dots, DB(A_n))$$

Selection refers to the manner in which AL/X chooses which, of many, questions is appropriate to ask next. (See Figure A-3). The manner in which this is done is important because human experts are characterized by their ability to focus quickly on the most efficient line of questioning. AL/X chooses and investigates that particular hypothesis which is most likely out of a group of perhaps several alternative hypotheses. The choice criterion is: Select that hypothesis which currently has the highest degree of belief. Once the hypothesis has been chosen, questions are selected which will speed resolution of the diagnosis. In AUDITOR, the process of selection consists of choosing which question to ask of the user, i.e., which question has the highest potential incremental weight impact on the hypothesis. The incremental weight of each of the remaining questions is calculated prospectively by the system. AL/X carries this out by scanning through all of the, as yet, unanswered questions, calculating the incremental weights which could



*IW = incremental weight

Figure A-3 Overview of the Method for Selecting the Next Question

arise from each question, as a result of a CV response by the user of either -5 or +5. (AND and NOT spaces are included in the calculation.) That question is chosen next which could conceivably cause the greatest impact on the strength of belief of the hypothesis. Since the potential impact of an AND space depends upon the user's answers to the component spaces, the order of questioning may vary from one consultation to the next depending upon the user's responses to the components of the AND spaces. This knowledge is utilized by the system in the form of rules often in the style IF:Evidence, THEN (to a specified extent):Hypothesis. The connectors AND, OR, and NOT may also be used.

The data are immediately incorporated into the system by the control and processed through the rule base. This action updates the rule base to reflect the user's knowledge and revises the degree of belief in the hypothesis which has been the focus of the expert's inquiry process. It is the degree of belief in the hypothesis (and its translation into probabilities) which constitutes the goal of the system and its expert judgment.

Goals with degree > 0.0 are:

The delinquent portion of this account should specifically be reserved for in the allowance for bad debts to a substantial extent (RESERVE). Prior degree was 0.0. Current degree is 13.8. At this point the goal is certain: Probability .95 or greater.

The above is AUDITOR's report that the user's answers to its questions have resulted in increasing the strength of belief of the truth of the hypothesis called RESERVE from 0.0 to 13.8 on the degree (of belief) scale. Prior to questioning of the user, RESERVE is presumed as likely to be true as false, that is, to have prior degree (or belief) of 0.0. A degree of belief (DB) or, simply, degree measures strength of belief in the goal hypothesis. The possible range is from -100 to 100. Since AUDITOR's report is based upon its Bayesian processing of subjective probabilities, the researchers utilized the report of degrees and probabilities merely as a guide to imputing a judgment from AUDITOR, and initially operated under the assumption that a probability of .90 (degree of 9.5) would indicate the need for an allowance.

A degree of zero indicates that the hypothesis is equally likely and unlikely, i.e., probability of 0.5. The mid-range of the degree scale from about -10 to 10 DB represents a large range of probability (about .10 to .90). Thus, small increments in degree may represent large changes in probability. This effect is considered desirable in this expert system since small changes in degree are considered to be

relatively more diagnostic when the current truth or falsity of the hypothesis lies in a doubtful or ambiguous area.

Only rarely, however, are the cue weights symmetrical. For example, the rule stated as, "This debtor is in bankruptcy-type proceedings" (BANKRUPT) carries $PW = 3.0$, $NW = -1.0$. These values indicate that to the experts the debtor's bankruptcy is considerably more diagnostic--reflected by the $PW = 3.0$ --of uncollectibility than the absence of bankruptcy is predictive of collectibility.

The reader may notice two consequences of these relationships. First, an apparently small portion of the available range on the DB scale, say from about -10 to 10, encompasses the rather large probability range .09 to .91. Additionally, since the effect of an evidential cue is additive upon the hypothesis, the diagnostic or inferential strength of any piece of evidence is at its highest when it impacts an hypothesis of neutral degree and decreases as more evidence is accumulated. This is believed to be a desirable feature in an expert system.

APPENDIX B:

ILLUSTRATION OF AUDITOR'S OPERATION

A hypothetical case situation is used to demonstrate the operation of the system. First, the facts of the situation, as they might be in the working papers, are described. Second, the consultative session between AUDITOR and the expert is shown, including annotations to highlight the features of AUDITOR. And, third, the results of the AUDITOR session are given in terms of degrees of belief (DB) and probability, which represent the revised likelihood of the hypothesis under scrutiny.

SITUATION

The delinquent amount owed by this customer, a regional, publicly held hardware chain, is \$82,000. This represents about two-thirds of the balance of one invoice approximately 75 days overdue on which the customer has recently paid \$40,000. In total, the customer owes \$320,000, the rest of it classified as current (45 days and under). The overdue amount, while large among delinquent accounts, is clearly not material to the client's financial statements. They show accounts receivable of \$20 million and current assets of \$100 million. The accounts receivable subsidiary ledger shows the most recent three-year history of the account, in which the customer has become overdue only one other time. He ultimately paid the full amount owed after an adjustment was made which represented about 2% of the invoice. Sales this year are above those of a year ago, probably reflecting a trend toward "do-it-yourself" repairs by hardware customers. Correspondence and other files indicate a long-term, on-going business relationship with no suggestion of any problems with the account.

The auditors did not attempt to confirm the account and there is no information available from an attorney or any other source outside of the client. The customer is not bankrupt. The client continues to make credit sales to this customer, who has paid other invoices within their due dates. According to the credit manager who has talked with them, the customer states emphatically he will not pay the balance which is overdue. The credit manager goes on to say, "Sometimes we give them a small adjustment when they put up a fuss like this, but they have been good customers for 15 years and they always pay." The "fuss" referred to in this case evidently resulted from a late delivery.

- CONSULTATIVE SESSION

The user first signs onto the system. AUDITOR then begins questioning the user about the facts of the case. The order of the questions depends on the user's responses. Material in brackets represents AUDITOR output, in parentheses the user's responses, and in asterisks the researchers annotations. The result of this session is that the

system determines that there is only .013 probability (-19.5 DB) that an allowance need be provided for this account. Other, similar accounts would go through a similar set of interactions.

APPENDIX C: NETWORK DESCRIPTION FILE

MODEL BADDEBT

VERSION ABDMA

SPACE RESERVE

TEXT DESCRIPTION

/* THE DELINQUENT PORTION OF THIS ACCOUNT SHOULD SPECIFICALLY BE RESERVED-FOR IN THE ALLOWANCE FOR BAD DEBTS TO A SUBSTANTIAL DEGREE */

INFERENCE

PRIOR 0.0

SPACE OUTSTAND

TEXT DESCRIPTION

/* WHILE SMALL PAYMENTS ARE BEING RECEIVED ON THIS ACCOUNT. THE OUTSTANDING DELINQUENT BALANCE IS GROWING LARGER */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 1 NW 0)

SPACE AVERAGE

TEXT DESCRIPTION

/* THE AVERAGE AGE OF THE UNCOLLECTED PORTIONS OF THIS ACCOUNT IS INCREASING */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 0.5 NW -0.5)

SPACE ALLBUTONE

TEXT DESCRIPTION

/* THIS CUSTOMER'S ACCOUNT BALANCE IS ALL CURRENT EXCEPT FOR ONE LARGE AND DELINQUENT CHARGE */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 0.5 NW 0)

SPACE ALBT1&NOTP

TEXT DESCRIPTION

/* ALLBUTONE AND NOTPAY */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (ALLBUTONE NOTPAY)

RULES CONSEQUENTS (RESERVE PW 2.0 NW 0)

SPACE CREDITSTOP

TEXT DESCRIPTION

/* CREDIT TO THIS CUSTOMER HAS BEEN STOPPED BY THE CLIENT */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 1 NW -0.5)

[...that this account has been assigned to a collection agency or lawyer? ?] (-5)

[...that this debtor has a good past record of ultimately paying substantially all delinquent balances? ?] (5)

[...that no payments have ever been received from this customer? ?] (-5)

[...that the customer has stated his intent to pay little or nothing of the delinquent balances? ?] (5) ' [...that despite the presence of this delinquent item, newer items have been fully paid? ?] (PG) *"PG" is the auditor's call to the system to present its conclusion. He feels either than an adequate amount of evidence, or all of the available evidence, has been entered into the system.*

[Investigated goals with degree of belief > -100.0 are: ...none at the moment.] *This should be read as indicating that the goal hypothesis has not been fully investigated. In other words, there are still questions unasked, but the user of the system prefers that they not be asked. Also, the system tells the user that it is set to report the degree for the hypothesis whenever it is over -100.0 .*

[Other goals with degree > -100 are:

The delinquent portion of this account should specifically be reserved-for in the allowance for bad debts to a substantial degree :RESERVE.

Prior degree was 0.0 , current degree is -19.5 . At this point this goal is probability $.05$ or less.] *The probability is only $.013$ (degree = -19.05 that the account will become uncollectible; therefore, the user will undoubtedly conclude no provision need be made for loss from this account.*

[How certain are you that despite the presence of this delinquent item, newer items have been fully paid? ?] (Q) * The system repeats the last question which is still unanswered, and the user tells the system to cease or Quit its questioning.*

[The current goal is whether or not the delinquent portion of this account should specifically be reserved for in the allowance for bad debts to a substantial degree. RESERVE. *The system is announcing its goal hypothesis.*

[How certain are you that your tests of cash receipts indicate that the delinquent portions of the account in question have been collected substantially in their entirety subsequent to the date of the balance sheet? ?] (-5) *On this line the system has given a "prompt" in the form of a question mark, and the auditor has replied with "-5" which indicates he is certain the account has not been substantially collected.*

[How certain are you that serious problems with the delinquent portions of this account were revealed through confirmation causing you to believe they are invalid?

?] (0) *No confirmation was requested, so the user responds with a "0" which means, in this case, the question is irrelevant.*

[...that legal counsel gives poor prospects of any significant recovery from this debtor? ?] (0) *There is no information available from a lawyer.*

[...that the confirmation request was returned by the postal service as undeliverable and the client for several months has been unable to communicate with the debtor? ?] (0) *The auditor responded to this question as irrelevant.*

[...that recent correspondence and other data in the customer's credit file supports your belief in the collectibility of this account? ?] (5) *The auditor believes the long-term business relationship will be continued by the customer, or that for other reasons the account will be collected.*

[...that recent collections toward the delinquent portions of this account are proceeding satisfactorily? ?] (5) *A substantial amount has recently been received.*

[...that the credit manager expresses a strong belief in the ultimate collection of substantially all this account? ?] (5)

[...that although a portion of this customer's total balance is still delinquent, he continues to be an active customer? ?] (5)

[...that this debtor is in bankruptcy-type proceedings? ?] (-5)

[...that economic factors are causing particularly bad effects on this customer's ability to pay? ?] (-5)

[...that the merits of this debtor's likely counterclaim against your client's suit indicate that a legal action would be fruitless? ?] (0) *The auditor has no information about a counterclaim.*

SPACE NRES&NTACT

TEXT DESCRIPTION

/* NORESPONSE AND NOT ACTIVE */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (NORESPONSE NOTACTIVE)

RULES CONSEQUENTS (RESERVE PW 2.0 NW -1.0)

SPACE NONCONTACT

TEXT DESCRIPTION

/* THE CONFIRMATION REQUEST WAS RETURNED BY THE POSTAL SERVICE AS UNDELIVERABLE AND THE CLIENT FOR SEVERAL MONTHS HAS BEEN UNABLE TO COMMUNICATE WITH THE DEBTOR */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 4.5 NW 0.0)

SPACE NCNT&NOTACT

TEXT DESCRIPTION

/* NONCONTACT AND NOTACTIVE */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (NONCONTACT NOTACTIVE)

RULES CONSEQUENTS (RESERVE PW 5.0 NW 0.0)

SPACE NOTPAY

TEXT DESCRIPTION

/* THE CUSTOMER HAS STATED HIS INTENT TO PAY LITTLE OR NOTHING OF THE DELINQUENT BALANCES */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 3.0 NW -1.0)

SPACE BANKRUPT

TEXT DESCRIPTION

/* THIS DEBTOR IS IN BANKRUPTCY-TYPE PROCEEDINGS */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 3.0 NW -1.0)

SPACE LAWYER

TEXT DESCRIPTION

/* LEGAL COUNSEL GIVES POOR PROSPECTS OF ANY SIGNIFICANT RECOVERY FROM THIS DEBTOR

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 5.0 NW -1.0)

SPACE BANK&RELAW

TEXT DESCRIPTION

/* BANKRUPT AND WRITEOFF AND LAWYER */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (BANKRUPT WRITEOFF LAWYER)

RULES CONSEQUENTS (RESERVE PW 9.0 NW -2.0)

SPACE ECONOMIC

TEXT DESCRIPTION

/* ECONOMIC FACTORS ARE CAUSING PARTICULARLY BAD EFFECTS ON THIS CUSTOMER'S ABILITY TO PAY */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 3.0 NW 0.0)

SPACE BANK&ECON&LAW

TEXT DESCRIPTION

/* BANKRUPT AND ECONOMIC AND LAWYER */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (BANKRUPT ECONOMIC LAWYER)

RULES CONSEQUENTS (RESERVE PW 7.0 NW -1.0)

SPACE GOODRECORD

TEXT DESCRIPTION

/* THIS DEBTOR HAS A GOOD PAST RECORD OF ULTIMATELY PAYING SUBSTANTIALLY ALL DELINQUENT BALANCES */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 2.0 NW 2.0)

SPACE LEGAL

TEXT DESCRIPTION

/* THE MERITS OF THIS DEBTOR'S LIKELY COUNTERCLAIM AGAINST YOUR CLIENT'S SUIT INDICATE THAT A LEGAL ACTION WOULD BE FRUITLESS */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 3.0 NW -1.0)

SPACE LEGAL&NPAY

TEXT DESCRIPTION

/* LEGAL AND NOTPAY */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (LEGAL NOTPAY)

RULES CONSEQUENTS (RESERVE PW 6.0 NW -1.0)

SPACE NPAY&GDREC

TEXT DESCRIPTION

/* NOTPAY AND GOODRECORD */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (NOTPAY GOODRECORD)

RULES CONSEQUENTS (RESERVE PW -2.0 NW 0.0)

STOP

SPACE NOTACTIVE

TEXT DESCRIPTION

/* NOT OF ACTIVE */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION NOT ACTIVE

SPACE CREDITMGR&ACT

TEXT DESCRIPTION

/* CREDITMGR AND (NOT PROBLEMS (NOT ACTIVE */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (CREDITMGR NOTPROBLEMS ACTIVE)

RULES CONSEQUENTS (RESERVE PW -3.0 NW 3.0)

SPACE CORRESPOND

TEXT DESCRIPTION

/* RECENT CORRESPONDENCE AND OTHER DATA IN THE CUSTOMER'S CREDIT FILE SUPPORTS BELIEF IN THE COLLECTABILITY OF THIS ACCOUNT */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW -2.0 NW 4.0)

SPACE CORRES&ACT

TEXT DESCRIPTION

/* CORRESPOND AND ACTIVE */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (CORRESPOND ACTIVE)

RULES CONSEQUENTS (RESERVE PW -3.0 NW 3.0)

SPACE CORRES&WORK

TEXT DESCRIPTION

/* CORRESPOND AND WORKOUT */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (CORRESPOND WORKOUT)

RULES CONSEQUENTS (RESERVE PW -3.0 NW 3.0)

SPACE WRITEOFF

TEXT DESCRIPTION

/* TOTAL WRITE OFF OF THIS ACCOUNT, IF REQUIRED, WILL REPRESENT A MATERIAL ADJUSTMENT */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 2.0 NW 0.0)

SPACE NORESPONSE

TEXT DESCRIPTION

/* THERE WAS NO RESPONSE TO YOUR CONFIRMATION REQUEST NOR TO A FOLLOW-UP REQUEST */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 1.0 NW -1.0)

SPACE COLAGENCY

TEXT DESCRIPTION

/* THIS ACCOUNT HAS BEEN ASSIGNED TO A COLLECTION AGENCY OR LAWYER */
INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 2 NW -1)

SPACE FORMEREMP

TEXT DESCRIPTION

/* THIS DELINQUENT ACCOUNT IS FROM A FORMER EMPLOYEE */
INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 2 NW 0)

SPACE NEWPAID

TEXT DESCRIPTION

/* DESPITE THE PRESENCE OF THIS DELINQUENT ITEM, NEWER ITEMS HAVE BEEN FULLY PAID
INFLUENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 1.5 NW 0)

SPACE NOPAYEVER

TEXT DESCRIPTION

/* NO PAYMENTS HAVE EVER BEEN RECEIVED FROM THIS CUSTOMER */
INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 2 NW -0.5)

SPACE ISSUENOTE

TEXT DESCRIPTION

/* THIS DEBTOR HAS ISSUED NOTES FOR THE UNPAID PORTIONS OF HIS ACCOUNT */
INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 0.0 NW -0.5)

SPACE OUTSTAVENP

TEXT DESCRIPTION

/* OUTSTAND AND AVERAGE AND NOTPAY */
INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (OUTSTAND AVERAGE NOTPAY)

RULES CONSEQUENTS (RESERVE PW 1 NW 0)

SPACE CREDITSTOP

TEXT DESCRIPTION

/* CREDITSTOP AND COLAGENCY AND NOPAYEVER */
INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (CREDITSTOP COLAGENCY NOPAYEVER)

RULES CONSEQUENTS (RESERVE PW 5 NW 0)

SPACE CREDITSTOP

TEXT DESCRIPTION

/* CREDITSTOP AND NOPAYEVER AND WRITEOFF */
INFERENCE

PRIOR 0.0

LOGICAL DEFINITION AND (CREDITSTOP NOPAYEVER WRITEOFF)

RULES CONSEQUENTS (RESERVE PW 3 NW 0)

SPACE COLLECTED

TEXT DESCRIPTION

/* YOUR TESTS OR CASH RECEIPTS INDICATE THAT THE DELINQUENT PORTIONS OF THE ACCOUNT IN QUESTION HAS BEEN COLLECTED SUBSTANTIALLY IN THEIR ENTIRETY SUBSEQUENT TO THE DATE OF THE BALANCE SHEET */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW -30.0 NW 1.0)

SPACE PROBLEMS

TEXT DESCRIPTION

/* SERIOUS PROBLEMS WITH THE DELINQUENT PORTIONS OF THIS ACCOUNT WERE REVEALED THROUGH CONFIRMATION CAUSING YOU TO BELIEVE THEY ARE INVALID */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 8.0 NW 0.0)

SPACE NOTPROBLEMS

TEXT DESCRIPTION

/* NOT OF PROBLEMS */

INFERENCE

PRIOR 0.0

LOGICAL DEFINITION NOT PROBLEMS

SPACE RIGOROUS

TEXT DESCRIPTION

/* COLLECTION EFFORT BEING APPLIED BY YOUR CLIENT TO COLLECT THIS ACCOUNT IS LESS RIGOROUS THAN IS DESIRABLE */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW 1.0 NW 0.0)

SPACE WORKOUT

TEXT DESCRIPTION

/* RECENT COLLECTIONS TOWARD THE DELINQUENT PORTIONS OF THIS ACCOUNT ARE PROCEEDING SATISFACTORILY */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW -3.0 NW 3.0)

SPACE CREDITMGR

TEXT DESCRIPTION

/* THE CREDIT MANAGER, OR OTHER COMPANY OFFICIAL, EXPRESSES A STRONG BELIEF IN THE ULTIMATE COLLECTION OF SUBSTANTIALLY ALL THIS ACCOUNT */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW -1.0 NW 2.0)

SPACE ACTIVE

TEXT DESCRIPTION

/* ALTHOUGH A PORTION OF THIS CUSTOMER'S TOTAL BALANCE IS STILL DELINQUENT, HE CONTINUES TO BE AN ACTIVE CUSTOMER */

INFERENCE

PRIOR 0.0

RULES CONSEQUENTS (RESERVE PW -2.0 NW 3.0)

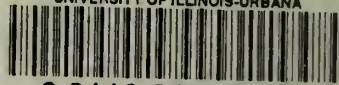
REFERENCES

- Abdel-Khalik, A. R., and El-Sheshai, K., "Information Choice and Utilization in an Experiment on Default Prediction," Journal of Accounting Research (Autumn 1980), pp. 325-342.
- Ashton, R. H., "A Descriptive Study of Information Evaluation," Journal of Accounting Research (Spring 1981) 19.
- _____, Human Information Processing in Accounting (American Accounting Association, 1982).
- Biggs, S. F., and Mock, T. J., "An Investigation of Auditor Decision Processes in the Evaluation of Internal Controls and Audit Scope Decisions," ROADS Paper No. 80-7, (April 28, 1980).
- Crosby, M. A., "Bayesian Statistics in Auditing: A Comparison of Probability Elicitation Techniques," Report 80-008, Center for Audit Research, University of Georgia (1980).
- Dawes, R. M., "The Mind, the Model, and the Task," in F. Restle, R. M. Shiffrin, N. J. Castellan, J. R. Lindman and D. B. Pisoni, eds., Cognitive Theory, Vol. 1 (Erlbaum, 1975), pp. 119-29.
- Duda, R., Gaschnig, J., and Hart, P., "Model Design in the Prospector Consultant System for Mineral Exploration," Expert Systems in the Micro-Electronic Age, D. Michie, Ed. (Edinburgh University Press, 1979).
- Dungan, C. W., "A Model of an Audit Judgment in the Form of An Expert System," Unpublished Ph.D. Dissertation, Department of Accountancy, University of Illinois (1983).
- Einhorn, H. J., "Expert Judgment: Some Necessary Conditions and an Example," Journal of Applied Psychology, 59 (1974), pp. 562-571.
- _____, Kleinmuntz, D. N., and Kleinmuntz, B., "Linear Regression and Process-tracing Models of Judgment," Psychological Review, 86 (1979), pp. 465-485.
- Ericsson, K. A., and Simon, H. A., "Verbal Reports as Data," Psychological Review, Vol. 87, No. 3, (May 1980).
- Ernst, G. W., and Newell, A., GPS: A Case Study in Generality and Problem Solving (Academic Press, 1969).
- Feigenbaum, E. A., "The Art of Artificial Intelligence: Themes and Case Studies of Knowledge Engineering," Proceedings of the 5th International Joint Conference on Artificial Intelligence, (1979).

- Harrell, A. M., "The Decision-Making Behavior of Air Force Officers and the Management Control Process," The Accounting Review, (October 1977), pp. 833-841.
- Joyce, E. J., "Expert Judgment in Audit Program Planning," Studies on Human Information Processing in Accounting, Supplement to the Journal of Accounting Research, (1976).
- Joyce, E. J., and Biddle, G. C., "Are Auditors' Judgments Sufficiently Regressive," Research Opportunities in Auditing Distribution Service, Peat, Marwick, Mitchell Foundation, N.Y. #80-4, (1980).
- Libby, R., Accounting and Human Information Processing (Prentice-Hall, 1981).
- Lichtenstein, S., Fischhoff, B., and Phillips, L. D., "Calibration of Probabilities: The State of the Art," in H. Jungermann and G. de Zeeuw (eds.), Decision Making and Change in Human Affairs, (Dordrecht-Holland: Riedel, 1977), pp. 275-324.
- Michaelson, R. H., "A Knowledge-Based System for Individual Income and Transfer Tax Planning," Unpublished Dissertation, Department of Accountancy, University of Illinois, (1982).
- Michie, D., "Knowledge-based Systems," Working Paper UIUCDCRS-R-80-1001, Department of Computer Science, University of Illinois at Urbana-Champaign, (1980).
- Newton, L. K., "The Risk Factor in Materiality Decisions," Accounting Review, (January 1977) 52, pp. 97-108.
- Nisbett, R. E., and Wilson, T. D., "Telling More Than We Can Know: Verbal Reports on Mental Processes," Psychological Review, Vol. 84, No. 3, (May 1977), pp. 231-259.
- Payne, J. W., Braustein, M. L., and Carroll, J. S., "Exploring Pre-Decisional Behavior: An Alternative Approach to Decision Research," Organizational Behavior and Human Performance, (February 1978) 22, pp. 17-44.
- Reiter, J., AL/X: An Expert System Using Plausible Inference, (Intelligent Terminals Ltd., 1980).
- Reitman, W. R., Cognition and Thought, (Wiley, 1965).
- Shields, M. D., "Some Effects of Information Load on Search Patterns Used to Analyze Performance Reports," Unpublished manuscript, University of North Carolina, Chapel Hill, (May 1980).
- Shortliffe, E. H., And Buchanan, B. G., "A Model of Inexact Reasoning in Medicine," Mathematical Bioscience, Vol. 23, (1975), pp. 351-379.

- Simon, H. A., "On How to Decide What To Do," Bell Journal of Economics, (Autumn 1978), pp. 494-507.
- Turing, A. M., "Computing Machinery and Intelligence," Mind, (October 1950), reprint in Creative Computing, Vol. 6, No. 1, (January 1980), pp. 44-53.
- Yu, V. L., Fagan, L. M., Wraith, S. et. al., "Anti-microbial Selection by a Computer: A Blinded Evaluation by Infectious Diseases Experts," Journal of the American Medical Association, Vol. 242, No. 12, (September 21, 1979), pp. 1279-1282.
- Zadeh, L., "A Theory of Approximate Reasoning," in Machine Intelligence, 9, J. R. Hayes, D. Michie, and L. Mikulich (Eds.), (Ellis Norwood Ltd., and John Wiley and Sons, 1979).

UNIVERSITY OF ILLINOIS-URBANA



3 0112 047450041